

中国政法大学

读 书 报 告

阅读书目：《Data-Driven Law》

题　　目：《数据时代下法律服务的动向、挑战与应对》

课　　程：法律服务与司法管理技术

姓　　名：王姝雯

学　　号：2001010118

学　　院：法学院

专　　业：法律职业伦理

研究方向：律师职业伦理

摘　要

数据因素的高活跃度使得法律数据有了良好的发展土壤，法律服务的复杂性和竞争性提升，法律服务范围由封闭转为开拓，除了传统的律师事务所外，更多新型的法律服务提供者出现，法律服务市场的活跃度有了很大的提升。法律数据挖掘由于其在法律数据和信息收集、储存、分析、预测等方面表现出的强大优势，深受一些法律服务提供者的青睐，但伴随的挑战与数据风险也亟待解决。

笔者的本篇读书报告在结合自身法律职业伦理专业的基础上，围绕本书第二章的“挖掘法律数据：采集与分析 21 世纪的黄金”进行深入分析，更加切合自身的研究旨趣与专业方向，希望能对本章的研究主旨产生更加精细化的体悟。

《数据驱动法律》一书囊括律师事务所的数据分析、法律数据挖掘、合同结解构、就业大数据转变、符号话语、法律服务质量量化、大数据揭示偏差等法律数据的新趋势，极具启发性。第二章提供了三个方面的研究旨趣，数据时代下法律服务的动向与应对机制，数据应用背景下法律职业伦理层面的挑战与风险规避以及法律数据发展中的 ABS、ALSP 与法律公司。对以上三个主题的分析与研究实则带来了很多启发与思考，有助于加深对当今数据时代下各种法律服务提供者在设立、运行、监管、评价等过程性认识，加深对新时代法律职业共同体的体系构建与新型法律服务提供者执业环境探索的思考。

关键词：法律数据挖掘；新型法律服务提供者；动向；法律职业伦理；挑战与应对

作者简介

本书书名为《Data-Driven Law》（《数据驱动法律》），由 Ed Walters 所著。本次的读书感想主要基于第二章——Mining Legal Data: Collecting and Analyzing 21st Century Gold（挖掘法律数据：收集和分析 21 世纪的黄金）展开。该分章的作者是 Kenneth A. Grady。

Kenneth A. Grady，作为法律技术领域的联合创始人，他带来了关于数据、自动化和人工智能在为合法消费者提供现代高效服务方面的作用的独特而及时的观点。他既是法律服务的企业买家，又是服务提供商，认为法律运营可以实现更高的效率和价值，尤其是对于企业买家而言。

任客座教授，作为学院 LegalRnD（法律创新中心）的一部分，在全球范围内教授多门课程，进行研究、出版和展示。课程和工作侧重于改善法律服务的提供以及对司法救助和实体法演变的相关影响。兴趣包括人工智能、律师-客户关系、技术、法律运营管理和法律的未来。

在律师事务所 Seyfarth Shaw LLP 兼职。此前，他是该公司 SeyfarthLean Consulting LLC 子公司的首席执行官。

担任 SeytLines.com 博客的编辑和主要作者。他的许多论文发表在美国国内外的法律期刊上。其行业观点在《华尔街日报》、《克莱恩斯》、《美国律师》和其他出版物中都有报道。曾是“全球 50 强”律师事务所的合伙人、三家财富 1000 强企业的总法律顾问、一家精品律师事务所的合伙人，并创立了一家独立律师事务所。曾在财富 1000 强公司的运营管理和行政部门担任高管职位。

14 年来，他一直是企业法律顾问协会的活跃成员，担任其董事会、价值挑战指导委员会和倡导委员会的成员。是 2017 年法律实践管理学院课程的候选者之一，被提名为 Fastcase 50，并因其在内部律师/外部律师关系方面的创新领导能力而受到英国《金融时报》的表彰。

原著第二章内容介绍

原著第二章“挖掘法律数据：收集和分析 21 世纪的黄金”。

在结构上，可以把本章分为八个部分。

第一部分“法律服务的复杂性”，描述了曾经法律服务提供者的知识具有地域封闭性，知识负担低。客户成本可控，更重视长期合作。现在，随着文件、案件、解释成文法的材料数量、司法判决以及州内部和各州之间的跨辖区法律服务增加，成文法范围拓展，法律服务比以往更加复杂，法律服务提供者获取知识的成本增加，负担加重，推动其改变工作结构。

第二部分“法律服务竞争”，阐述了数据时代给过去和现在的法律服务主体带来的新变化。在过去，数据运用不那么丰富，有法律服务需求的主体选择法律服务提供者主要基于地理位置、获取法律信息的途径、法律服务提供者的谱系、市场声誉、成本等因素做出选择。而在数据和技术广泛应用的现在，人们选择法律服务提供者主要基于成本因素，更多的可能涉及品牌因素。以律师事务所的品牌为例，这是律所对外宣传的招牌，律所文化、业绩、网站、律师资质、办公环境等都对律所业务的拓展起到重要影响。另外，在数据愈发活跃的当下，此前律所之间相互竞争的单一局面被打破，涌入一批新的组织（法律服务咨询公司等替代法律服务提供者）。

第三部分“仅基于经验的建议与基于数据挖掘和经验的建议”，系统 1 即“仅基于经验的建议”，是指纯粹的“根据我的经验”，这种模式非常快速，具有启发式，但多适用于单个法律服务提供者进行较小样本量的处理，且个人的经验无法避免认知偏差的存在。系统 2 即“基于数据挖掘的建议”，通过处理和深入分析数据来生成答案、增加价值，对数据的依赖性极高，比系统 1 慢，不能同时处理多项任务，且单采系统 2 则很难与世界互动。故在此采用综合性的方式，当数据难以收集和分析时应用系统 1；如今，当数据更容易收集和分析时，将系统 1（经验）与系统 2（数据）结合进行法律数据的分析和建议的提供，更得人意。

第四部分“数据时代与法律服务风险”，由 19 世纪的黄金时代，到 20 世纪的石油时代，再到 21 世纪的数据时代，每个世纪都有每个世纪的标志。互联网一方面给了法律服务以新的数据采集工具，提供了更丰富的答案和分析途径；另一方面，在数据的运用中，法律服务的风险也随之上升。知识风险产生于法律服务提供者利用其所知或应知来提供法律服务之时，这种风险由于尚处在法律服务提供者的可控范围内，故而显得比较平稳。但无知风险呈现出不可控的

增长状态，源于法律服务提供者未能收集和使用更准确地反映现实世界的数据。数据带来的解决方案集可能性的不断增加更使得无知风险增加。

第五部分“将数据挖掘与提供法律服务整合”，在数据时代带来分析工具便利化的机遇下，如何把握好数据挖掘的红利，并将其与法律服务的提供高效整合，是值得探讨的问题。提示法律服务提供者关注到法律数据挖掘所得的各数据范围，避免细节信息的遗漏。

第六部分“数据挖掘程序的需求和组件——数据管理计划（DMP）”，该部分围绕 DMP 展开，以数据生命周期的八个组成部分为重点，即计划、收集、确信、摘要、保存、发现、完整和分析，共同构成了法律数据挖掘的全过程。

第七部分“数据的可移植性与文本数据的价值”，描述了用户对数据占据主动权的今天，数据是可移植的。运用数据系统，法律服务提供者可以从数字文本文件的大量材料中提取数据，通过对文本应用结构来解析文档。

第八部分“新兴数据源的挑战与 21 世纪数据挖掘在法律服务提供中的重要性”。带有嵌入式传感器的产品数量增加，将数据流发送给任何被授权接收的人，者带来一个重要问题，如何对收集到的客户法律数据履行妥善的保密义务。专业法律服务提供者通过运用软件对专有数据、公开可用的数据和客户数据进行收集分析，最终形成向客户提供的服务。这是该部分的核心公式所在，也是法律服务提供者进行法律数据挖掘的重要流程。

一、数据时代下法律服务的动向与应对机制

数据时代下，法律服务相较于传统模式而言呈现出新动向，特别是法律服务如何与数据挖掘结合。在传统的法律服务提供者如律师事务所、法院、检察院等之外，出现了新型法律服务组织如法律服务咨询公司、ABS、ALSP 等。针对这种动向，法律服务主体的应对机制如何，如何达成两者间的协调与合作，是当今数据时代下高效收集和分析法律数据的重要内容。

（一）法律服务的复杂性与竞争性

在过去，法律服务活跃范围狭窄，活跃度不高，法律服务提供者的知识具有地域封闭性，知识负担在广度层面较低。客户和法律服务提供者均更倾向于长期合作，客户往往选择同一地域内更加熟悉、可靠的法律服务提供者，在这一阶段，客户成本是可控的。而随着科技的敏感度不断上升，数据的活跃程度大大提高，使得法律活动中的文件、案件、解释成文法的材料数量和司法判例增加，成文法范围不断扩展，原来区域封闭性的法律服务壁垒逐渐被打破，拓宽至跨区域甚至跨国范围。多重因素的拓展使得法律服务比此前变得更加丰富和复杂，其中涉及的法律知识、行业习惯、标准规程不断丰富，对法律服务提供者提出了更加严格的要求。

随着法律活跃度的提升，法律服务提供者的主体数量和类型也呈增长态势，法律服务的竞争性加强。除了典型的公共法律服务提供者，如人民法院、人民检察院等之外，私法律服务提供者和组织中也出现了新动向。律师事务所由原来的相互竞争发展到还与其他新型法律服务组织竞争，如法律服务咨询公司、替代法律服务提供者（ALSP）、替代性商业结构（ABS）甚至外国法律服务提供者。其中，法律服务提供者和法律服务组织的范围又有所不同，前者是指为委托人提供法律服务的律师和无执业资格的个人，它避免了贬义词“非律师”以及解析哪些服务构成“未经授权的法律实践”这一日益无关紧要的任务。后者包括作为法律服务供应链一部分的所有实体。[1]

（二）法律数据挖掘

法律数据挖掘，简而言之即为收集和分析法律数据的全过程，主要包括法律服务提供者提供服务时收集和创建材料的结构、内容以及提供法律服务的流程。在对法律数据进行定义后，进入到对数据的预处理阶段，从原始法律数据中抽取正确、可靠且所需的数据，进入到数据挖掘的内核即挖掘算法中，从众多模式集合中筛选出符合需要的表达与解释类型，通过分析得出最后的数据，传达到用户手中。

“传统的在线法律服务主要分为法律法规案例数据库模式和律师个人营销推广模式(例

[1] See Ed Walters, *Data-Driven Law*, Boca Raton: CRC Press, 2019, p.50.

如在线解答咨询),它们都没能利用海量数据本身进行价值发掘。”[2] 在新的数据时代下，对于第一种模式，可以通过数据分析和挖掘过往相关案例的法律法规及规范性文件的适用、判决书的逻辑、类似案件个性化层面之外的普适性裁判规则以及某类案件的总体胜负率。将由此形成的法律分析产品向律师事务所出售，使律师事务所和律师在成为平台和系统的使用者的同时，也是某些特定内容的提供者，并且这种分析方式也有助于加强对公权力司法机关进行监督，推动公共法律服务透明化、私人法律服务高效化和便利化。第二种模式，根据目标客户的法律需求，在系统中分析并向其推荐契合的律师为其提供法律服务，缩短需求时间，促进法律资源的流动，为客户提供符合期望的委托选择。数据挖掘技术可以发现海量数据中潜在的、人工容易遗漏收集和分析的数据信息，可以在“数据爆炸”的困境中运用系统和数据平台进行有条不紊的收集、抓取和分析，把握高质量数据。[3]

（三）传统法律服务提供者的应对机制

在数据喷发的时代，传统法律服务提供者在新的数据环境下如何调整自身，如何应对新型法律服务组织的挑战与竞争，是值得探究的问题。

1. 公共法律服务提供者

“2013 年 8 月，最高人民法院院长周强在最高院全国第四次司法统计工作会议中，第一次提出‘大数据、大格局、大服务’的理念，从此拉开了我国司法大数据的序幕。”[4] 中国裁判文书网、中国庭审公开网的建设，将案件庭审实况和裁判文书置于网络数据的大环境监督下，更加公开以受监督，有利于增强我国司法的公信力，及时纠错。法律数据流由分散式、零星式变得更加聚合、丰富，裁判文书、审判信息等法律核心数据的共享机制由确立到完善，司法信息化建设强力推进。

在数据化、信息化公共法律服务的建设中，尤以智慧法院的建设为示范内容。“2016 年《国家信息化发展战略纲要》将建设‘智慧法院’列入国家信息化发展的战略中。同年《“十三五”国家信息化规划》中明确指出：支持‘智慧法院’建设，推行电子诉讼，建设完善公正司法信息化工程。”[5] 从纯粹人工的案卷整理、信息归档到案件受理、审判、执行、监督等各环节信息化和人工处理的综合，从案卷的书面化、庭审的现场化到电子式案卷信息的记录、互联网庭审的公开同步直播和案件办理流程的数据可视化，司法信息得到更高程度的公开和监督，司法公平正义有了更加稳定的保障。在机器学习层面，首先需要对输入文件的规范性格式进行严格要求，保证其能够被法律分析系统和平台准确对应且识别，继而进行后续的数据采集和分析。这种大数据分析方法在审判辅助办案系统中多用于数据抓取、追溯、监

[2] 胡凌：“大数据兴起对法律实践与理论研究的影响”，载《新疆师范大学学报(哲学社会科学版)》2015 年第 4 期，第 108~113 页。

[3] See Zhongping Xu,Xu Zhongping,Zhang Tao,Zheng Rongrong,Li Xiangna,Li Dong, “Knowledge Discovery and Analysis of ICT Customer Service Data Relying on Multi-frequency Mode Data Mining”，Journal of Physics: Conference Series，4（2020）.

[4] 梁增辉：“大数据时代下的法律服务新模式”，载北京市律师协会、云南省律师协会、贵州省律师协会、四川省律师协会：《第三届“京云贵川”律师实务研讨会论文集》，2019 年版，第 6 页。

[5] 秦永彬等：“‘智慧法院’数据融合分析与集成应用”，载《大数据》2019 年第 3 期，第 36 页。

控、预测和总结上。此外，在公共法律服务提供者运用的法律数据分析工具上，根据需求的不同，有不同的工具选择。“运用 Python 等编辑的数据抓取工具，适合于在裁判文书网等网站自动抓取数据信息，且可以进行一定程度的词汇分析，是较为灵活的数据采集方式，但需要一定的编程能力。”[6] Excel 表格是传统的整理、记录和收集工具；SPSS 统计分析工具在进行数据统计和分析应用中较为广泛，但其必须使用规范化、结构化的格式进行录入，以便于系统识别。检察院的数据建设也随着智慧法院的深入发展投入进程。数据时代下的公共法律服务有效突破了司法行政改革的固有路径，利用信息技术和数据系统平台开放数据端口，实现法律信息的互享互通，使多种法律服务融合到多端口接续的“一站式”服务中，为人民群众提供便利。[7]

2. 私法律服务提供者

数据要素越活越，由此导致的私法律服务提供者的差距会越来越大。精明的私法律服务提供者占据更加完备的数据资源，其建立或使用的法律数据系统、平台和技术更加成熟，跨领域的数据资源更加丰富，其有更多的资本、团队来投入建设，但相较之下的小型律所由于缺乏这方面的资源，让其独立构建法律数据系统或平台十分困难，特别是在当今数据发展态势飞速上升的境遇下，小型律所仅凭人力将愈发难以生存，最终解散或被大型律所合并。

客户是律所业务拓展中的重要一环。在数据要素不活跃的时段，律所的业务范围较为闭塞，聚焦内部工作流程，而当数据跨区域快速流动时，律所更多地关注到客户、案源拓展、法律数据、律师评价中去。运用离群数据的分析方法，可以高效地找寻目标客户。离群数据简单来说就是把相似的东西归为一组。 比如对客户进行分组时，关注多个衡量因素如业务范围符合度、标的额等，以制定精确的策略。首先，获取所需的客户数据并执行简单的数据清洗，筛选出重要的数据与律所及具体律师进行匹配；其次，根据客户的特征构建多维度的特征向量；最后，检查输入的数据是否符合结构化标准，继而构建模型以搜索目标客户。反之，在离群目标客户选择的另一面，也有客户、律所、律师监管机构等对律师评价体系的构建。律师职业本质上还是服务业，其提供的是一种具有公共性的法律服务，对服务质量的评价对本单法律服务的提供和后续客户的选择上具有重要的参考价值，所以对律师行业评价体系的构建是极具现实意义的。虽然行业内已有专门针对律师的专业评级制度，但该制度的适用范围较窄，已经无法满足逐渐上升的多方面评价需求，因此亟需建立大数据下的律师评价体系，包括权威的律师身份认证体系、律师业务范围体系、客户评价体系、过往案件代理评价体系等，以达到直观的监管和对律师勤勉规范执业的督促。

私法律提供者在数据时代下，从以往的纯人工式法律信息收集过度到互联网平台、系统的法律数据挖掘和个人经验结合的信息收集分析，不仅在于时代的大背景，也在于数据在法

[6] 杨凯：“公共法律服务智能应用新视野——以人工智能技术与审判辅助办案机制建构为中心”，载《湖北警官学院学报》2020 年第 5 版，第 30 页。

[7] See Gong Yan, “System Construction and Practice of ‘Internet +’ Public Legal Service in the New Era”，Modern Management，5(2019) : 634~642.

律领域的应用的确能给私法律服务带来便利和高效率。一般情况下，人的记忆力和专注度是有限制的，在案件错综复杂、法律各方面的要素过多时，采用结构化的数据输入能够保障法律数据分析时的清晰条理，减少统计和分析的失误。特别表现在检索层面，对于法条、案例、文书的检索和运用也更加全面和便捷。在法律文书层面，“利用人工智能技术，通过机器识别账本、票据、合同文本等,将纸质档案转换成电子文档，借助大数据全景式记录的信息，方便律师为客户代写法律文书、提出法律意见”[8]。通过收集、检索和应用法律数据，在自己熟悉的案件领域外，律师还可以对重大疑难复杂的案件有所突破，通过对过往类似案件的梳理把握宏观规律，结合个案的具体条件更高效地熟悉和分析案件，更好地服务于当事人。大数据技术的运用还能够突破样本采集方面的局限，极大扩展样本采集的范围，其中所隐含的信息能够帮助律师突破思维盲区，推动其思维从演绎推理转变为归纳推理和类比推理，从依赖因果关系转变为寻找相关关系，在预测案件方面起到很大的贡献。[9] 在律所和律师提供的基础法律服务之外，数据收集、检索、存储分析等功能带来的新型法律服务也进入视野。大型律师事务所背后有强大的资金、资源支持，为这类新型法律服务的开展提供了坚实的后备基础。

3. 法律服务系统和平台

法律服务系统和平台，是数据时代下解放一定人力的重要途径。常见的法律服务系统及平台如北大法宝、中国裁判文书网、中国庭审公开网等，但这些为人所熟知的平台更多的集中于检索功能，少有数据分析、预测、过程性服务、法律资源整合调度等新型功能。创新型的互联网法律服务系统或平台与客户的关系其实类似于生产者、销售者与消费者的关系，涵盖对于目标客户与服务提供者的需求匹配、过程性评价保障、平台或系统运行中参数的更新和技术创新等内容。

在法律服务系统运作的过程中，对于法律结构化文本的挖掘是核心内容，这是获取法律信息的第一也是最重要的一步。通过适合个案的对应算法，如 C4.5[10]、k-means[11]、Apriori[12]、EM[13]、PageRank[14]、Adaboost[15]算法等，挖掘所得法律文件中信息的全部价值（至少是所需价值），

---

[8] 曹远军：“大数据应用背景下的法律服务探索”，载《中国律师》2020 年第 6 期，第 70~71 页。

[9] 参见张明新、李冠叶：“大数据时代我国法律服务行业的新发展”，载《江苏师范大学学报(哲学社会科学版)》2020 年第 6 版，第 93~103 页。

[10] C4.5 分类树是决策树算法中最流行的一种，它是一种监管学习，给定一堆样本，每个样本都有一组属性和一个类别，这些类别是事先确定的，通过学习得到一个分类器，该分类器能够对新出现的对象给出正确的分类。

[11] k-means 算法，也被称为 k-平均或 k-均值，是一种得到最广泛使用的聚类算法。它将各个聚类子集内的所有数据样本的均值作为该聚类的代表点，主要思想是通过迭代过程把数据集划分为不同的类别，使得评价聚类性能的准则函数达到最优，从而使生成的每个聚类内紧凑，类间独立。这一算法不适合处理离散型属性，但对连续型具有较好的聚类效果。

[12] Apriori 算法是一种最有影响力的挖掘布尔关联规则的频繁项集的算法。该算法过程分为两个步骤：第一步通过迭代，检索出事务数据库中的所有频繁项集，即支持度不低于用户设定的阈值的项集；第二步利用频繁项集构造出满足用户最小信任度的规则。

[13] 最大期望（EM）算法是在概率（probabilistic）模型中寻找参数最大似然估计或者最大后验估计的算法，其中概率模型依赖于无法观测的隐藏变量（Latent Variable）。常用于机器学习和计算机视觉的数据聚类（Data Clustering）领域。

[14] PageRank，网页排名，又称网页级别，是 Google 专有的算法，用于衡量特定网页相对于搜索引擎索引中的其他网页而言的重要程度，体现网页的相关性和重要性，在搜索引擎优化操作中是经常被用来评估网页优化的成效因素之一。

[15] Adaboost 是一种迭代算法，其核心思想是针对同一个训练集训练不同的分类器（弱分类器），然后把这些弱分类器集合起来，构成一个更强的最终分类器（强分类器）。Adaboost 算法本身是通过改变数据分布来实现的，它根据每次训练集之

正是一步步的挖掘构成了对法律文本最终的透彻分析。另外，文本挖掘技术克服了过去手动关键词搜索的限制，剔除了太多不相关案例记录的误报，以自动化该过程，克服现有方法中的困难。[16]

互联网法律服务平台的准入，是法律数据建设中的重要一环，这决定了客户所接触的法律服务提供者的质量，继而影响整个法律服务提供进程的质量和评价。故要明确平台关于律所、律师和客户应遵守的行为规范，严格律所和律师准入平台的资质审查，载明律师的执业资质、擅长的业务领域，对曾受处分、违法违规的律师不予准入，此为事前准入。在过程性评价和监督方面，在保障个人信息安全的情况下，对每一位客户和执业律师的不涉密行为进行宏观记录和监督，建立个性化的诚信档案，使客户、律师和律所在法律服务的获得和提供中的信息基本对称，以示公平，防范诚信风险。另外，通过在法律服务平台中法律设区交际圈的完善，促进律师、律所与客户之间即时的交流沟通，增进互信，增加客户对互联网法律服务平台的粘性。[17] 以法小宝为例，该智能问答产品提供涵盖法律咨询、法律知识查询、费用计算、文书模板、人工咨询[18]等在内的多项内容。在法律咨询和人工咨询方面，通过区分案件类型、设置多渠道，进行对案件的分流，初步的智能问答解决平台自身能够回答的简易问题，过滤出的高级问题再转接至人工，高效配置法律资源，减轻繁琐的人工负累；在文书模板方面，提供包括起诉状、答辩状、代理词、合同等多种法律文书的基本格式，以供使用者参考；在费用计算方面，通过对于结构化文本的挖掘，自动识别并对应各项费用，再根据系统内机器学习的各项费用的计算标准予以核算。“第三方法律服务平台运用数据分析统计的方式，采集供需双方的用户信息，根据用户的关注领域、使用习惯、律师的专业方向等信息进行整合，将高度相关的律师和客户互相推送，从而实现利用数据寻找目标的高效匹配模式。”[19]

（四）律师、律所与新型法律服务提供者的业务界限

数据发展给法律服务带来的竞争性使得传统的律师事务所面临新型法律组织的挑战，如法律服务咨询公司、ABS 替代性商业结构、ALSP 替代法律服务提供者(如法律研究服务)，甚至外国法律服务提供者。法律服务提供将从主要围绕内部工作流程的模式转变为围绕客户结果的模式。律师事务所将完成从基于服务的按小时计费模式向更接近产品订阅模式的演变，带来增长、高利润率的可持续经常性收入和新的更高价值的服务机会。[20]律师事务所坐拥大量数据，但目前需要思考如何面对数据发展的形势作出应对和运用。数据与技术结合的形式将

中每个样本的分类是否正确，以及上次的总体分类的准确率，来确定每个样本的权值。将修改过权值的新数据集送给下层分类器进行训练，最后将每次得到的分类器最后融合起来，作为最后的决策分类器。

[16] See M.F.M Firdhous，“Automating Legal Research through Data Mining”, International Journal of Advanced Computer Sciences andApplications, 6(2011).

[17] 宋皓：“着力破解‘互联网+法律服务’发展的四大困境”，载《人民论坛》2019 年第 28 期，第 94~95 页。

[18] 参见“大数据 AI 助力法律服务——法小宝·智能问答”，载《中国律师》2019 年第 12 期，第 86 页。

[19] 张明新、李冠叶：“大数据时代我国法律服务行业的新发展”，载《江苏师范大学学报(哲学社会科学版)》2020 年第 6 版，第 93~103 页。

[20] See Simon Drane, “How Data Will Enable the Shift Towards the Productisation of Legal Services”, Legal Information Management, 3(2019): 138~142.

使新型法律服务提供者能够提供新形式的产品化产品，推动新的经常性收入和更高价值的商业咨询服务。

传统律师事务所和律师在面对新型法律服务提供者时，是否有必要对自身进行适时的优化，律所、律师与新型法律服务提供者在业务提供方面存在何种业务交叉的范围和界限，如何保证律师公共性与公司营利性在业务的提供、合作中不产生冲突并达到高效融合，是值得探讨且需要在实践中予以明确的。律师在所属律所执业过程中，不能以律师身份与新型法律服务提供者如法律服务咨询公司订立劳动合同，但可以非律师身份与其签订劳动合同，这是《劳动法》、《劳动合同法》等民事法律中对于不禁止双重劳动关系的要求。以法律服务咨询公司为例，在律所执业的律师虽不能以律师身份与法律服务咨询公司签订劳动合同，但可担任其法律顾问，为其日常业务涉及的法律咨询、案件代理等提供业务范围内的服务。这种律师执业限制的规定主要基于《律师法》第 10 条[21]和第 12 条[22]中对于律师职业场所和兼职律师的限制。律师的执业场所仅限定为一个律师事务所，不得在多个场所同时执业。另外，兼职律师较为特殊，其在律师身份之外，还作为高等院校、科研机构中从事法学教育、研究工作的人员，并符合《律师法》第 5 条[23]的规定，这仅为律师具有双重职业身份的唯一例外。律所、律师与新型法律服务组织业务界限的重合还在于三者间合作机制的建立，新型法律服务组织由于可能涉及更多的商业、金融、数据、医药等业务领域，其接触的案源需求和范围较广，可为律所和律师进行案源的介绍，由律所和律师代理案件，合作共赢。律师可以设立法律服务咨询公司，但基于其公共性的职业性质，不能担任该法律服务咨询公司的股东或法定代表人（涉及到一定的营利性，会与公共性产生冲突），且基于同一执业场所的规定，其也不能在其中以律师身份执业。

二、数据应用背景下法律职业伦理层面的挑战与风险规避

数据时代下，信息的活跃度增强，分布范围扩展，信息的高度自由带来了许多法律职业伦理层面的挑战与风险，如何进行挑战的应对与风险的规避，是此处探讨的重要内容。

（一）对客户数据的保密义务、隐私泄露风险及应对

[21] 《中华人民共和国律师法》第 10 条规定：“律师只能在一个律师事务所执业。律师变更执业机构的，应当申请换发律师执业证书。
律师执业不受地域限制。”

[22] 《中华人民共和国律师法》第 12 条规定：“高等院校、科研机构中从事法学教育、研究工作的人员，符合本法第五条规定条件的，经所在单位同意，依照本法第六条规定的程序，可以申请兼职律师执业。”

[23] 《中华人民共和国律师法》第 5 条规定：“申请律师执业，应当具备下列条件：
（一）拥护中华人民共和国宪法；
（二）通过国家统一司法考试；
（三）在律师事务所实习满一年；
（四）品行良好。
实行国家统一司法考试前取得的律师资格凭证，在申请律师执业时，与国家统一司法考试合格证书具有同等效力。”

律师保密义务[24]是指律师对执业时接触到的委托人数据、隐私或商业秘密，未经委托人同意不得向第三人公开。在法律数据的传输、储存、修改等过程中，大量法律数据流涌入、交汇，需要对这些数据流进行有序梳理和储存。在客户元数据方面，其支持如指示存储位置、历史数据、资源查找、文件记录等功能，条理性强，形成了对于各个客户法律数据和信息的电子化更迭的动态记录。但这一动态的数据记录中必然包含了大量带有各个客户隐性数据标签的法律数据，对其进行分类加密储存，规避客户隐私泄露风险，是法律职业伦理层面的一个挑战。

数据挖掘技术允许利用大量数据来发现以前未知的、可能有用的知识，以及数据的隐藏信息和价值，信息技术发展的高速和广度更是增加了数据挖掘的力量和潜力。然而，个人敏感信息的隐私在这个过程中普遍没有得到尊重，这在一些应用中产生了伦理和法律问题。[25] 网络挖掘技术的一部分旨在分析用户的行为，以不断改进访问网站的结构和内容。但使用网络挖掘提供的强大处理工具可能会威胁到用户的隐私。当前关于隐私问题的法律研究提出了一种基于网络挖掘独特属性的综合方法，可以在每个特定背景下确定可能威胁个人隐私的行为，以确定哪些技术和用途有害于客户隐私。[26]但我们如何对客户隐私作出界定，以法律数据挖掘为例，委托人的隐私包括其个人信息和资料、委托事项的具体内容，涉及案件的还包括案件进程信息等，公司型客户还涉及商业秘密，综合来看，客户隐私的范围还是较广的，对其承担的保密义务也十分重要。

在客户数据的保密和隐私泄露风险的应对模式上，可以作出这样的尝试：①在收集数据前做出收集标准和隐私声明。挖掘者应当在收集数据前就制定出相应标准，阐明用户的哪些数据和信息会被收集，以供客户选择是否要进行此次委托，或者能否根据个案和个别客户性质的特殊性而对这些收集的数据范围作出调整。确定了数据收集范围后，采集到的数据在此后的进程中将被用于哪些可能的领域或用作何种用途，均需要明确告知客户，进一步的细节还可与之沟通；②相应的标准或规程制定后，进入到数据收集的过程性阶段，法律服务提供者应将后续对于客户数据的分析和处理过程对客户明示，提升透明度，给予客户保障与信赖；③法律服务提供者对收集到的数据要妥善保管，践行保密义务，在不违反国家法律法规规定的情况下防止第三方获取，造成用户信息泄露。“当数据源中含有客户身份、健康状况、财产、信用等方面的信息时,Web 挖掘软件应考虑给每个客户分配一个经过加密处理的标识。系统的输入应为匿名数据,而匿名数据向特定客户的转换机制应当只有 Web 站点才知晓;当数据源中含有客户具体数据,以至于该客户可以被惟一地辨识出来时,Web 挖掘系统应尽量避免采用客户单个的具体数据,而是采用客户的汇总数据,即只对客户汇总数据进行挖掘。另外,Web

---

24 《中华人民共和国律师法》第 38 条规定：“律师应当保守在执业活动中知悉的国家秘密、商业秘密，不得泄露当事人的隐私。

律师对在执业活动中知悉的委托人和其他人不愿泄露的有关情况和信息，应当予以保密。但是，委托人或者其他人准备或者正在实施危害国家安全、公共安全以及严重危害他人人身安全的犯罪事实和信息除外。”

25 Mehmet Cudi Okur, “On Ethical and Legal Aspects of Data Mining”, Journal of Yasar University, 11(2018).

26 Juan D. Velásquez, “Web mining and privacy concerns: Some important legal issues to be consider before applying any data and information extraction technique in web-based environments”, Expert Systems With Applications, 13(2013).

站点应建立一个隐私控制的框架,在客户隐私数据和其他数据之间构筑起一道防火墙。建立一种机制,使之能够对数据仓库逻辑模型中所有与用户相关的数据进行确认和过滤,允许在Web挖掘之后对结果进行检查。”[27] 这些隐私处理从源头和过程中对客户隐私泄露的风险进行规避，履行保密义务。另外，近日有关部门根据《中华人民共和国数据安全法》[28]等相关法律法规，陆续对网络数据公司开展了窗口指导工作。例如，聚法案例网站去除了裁判文书部分当事人的个人信息，以及案号、审理法院、审理法官等信息，主文其他内容不受影响，这也是对法律数据信息的隐私处理，只不过对象并非律所或新型法律服务提供者的客户，而是对数据系统或平台中第三方隐私的保护。

（二）律所管理视角下打破数据联通壁垒的尝试

数据时代下，对法律数据的收集、分析和运用在法律职业共同体的执业过程中具有重要价值。在特定的法律职业共同体内构建法律数据互联互通的路径，需要根据实际情况打破必要的数据联通壁垒，开放数据交互接口。“目前的法律科技服务需要建立一个体系，例如打造一些包罗国内外法律诉讼案例的数据库，推动国内外律所、律师和法律资源的交流互通。”[29] 律师事务所内部也需要搭建数据互联互通的平台和模式，如各律所内部形成的各种合同模板、同级不同小组之间能对数据达到多大程度的互通等等。出于对客户隐私的保密义务要求和律所内部信息互通互享的平衡，可以采用这样一种模式：同级的各项目组成员之间对于各自负责的客户数据是知晓的，但对其他项目组负责的客户数据是保密状态，可提交至共同的上级进行数据的加密化分析处理，从中提炼出有价值的、可不涉密的标准化信息实现各项目组的共享，在共享过程中只是接触到标准化、匿名化的有参考价值的模板化数据，并不涉及对具体显名客户数据和隐私的侵犯。由此进行学习借鉴，提升律所的工作效率，实现律所、律师和客户的共赢。

（三）法律数据系统和平台的维护和推广

在对法律数据系统和平台的建设中，对其的维护和推广是非常重要的过程。改正法律数据软件系统在使用过程中发现的隐含错误，扩充用户提出的新的功能及性能要求。在数据维护方面，需要对客户数据的更改进行及时更新，做好数据的储存、备份和恢复，定期检查法律数据库的运行状态是否正常，保证整体的数据平稳运行，提供给用户良好的体验。类似于律师的业务推广规则，律所或新型法律服务组织在对法律数据系统进行推广时也需遵守一定的规则。特定的律所拥有了自己的法律数据系统或平台后，继而通过关联规则、离群数据寻找目标客户，进行有效推广，扩大受众面，挖掘潜在客户，这也需要系统本身的精密及与客户的匹配度。另外，在实践过程中需要减轻法律数据系统和平台对特定主体的依赖程度，打

[27] 曹玉平：“数据挖掘中的信息伦理问题研究”，载《情报理论与实践》2008 年第 5 期，第 695~698 页。
[28] 2021 年 6 月 10 日，第十三届全国人民代表大会常务委员会第二十九次会议通过《中华人民共和国数据安全法》，自 2021 年 9 月 1 日起施行。《中华人民共和国数据安全法》是为了规范数据处理活动，保障数据安全，促进数据开发利用，保护个人、组织的合法权益，维护国家主权、安全和发展利益，制定的法律。
[29] 金凤：“与科技联姻，法律服务体验或被重构”，载《科技日报》2021 年 6 月 29 日。

造可传承的、更加独立的运行系统，以防止空缺和系统瘫痪。

三、法律数据发展中的 ABS、ALSP 与法律公司

相对于传统律所的发展模式而言，数据时代下新型法律服务提供者的出现带来一些新的发展机会。ABS、ALSP 以及法律公司均是具有竞争力的法律数据产物，这些模式在国外发展如何，以及与中国的本土化发展形势有多大的适配度，将在本章予以探讨。

（一）ABS

律师作为一种公共性的职业，必须与一般的以营利为目的的商业活动划清界限。但不可否认的是，律师的业务的确可以通过法务市场中经济、资源、数据、信息等因素的活跃来得到改善。但如果法务市场中的市场原理得到充分贯彻甚至市场分子过于活跃，律师会不会在与非专业的法律服务人员的竞争中失去其身份特性，或者说律师在这种浮躁的经济社会中丧失原本公共性的本心，从而偏入逐利的轨道，这就是法务市场中极其深刻的二律背反。特别是西方发达国家目前盛行的 ABS（Alternative Business Structure）替代性商业结构、法律服务零售店面（如 Axess Law[30]）等，这种法律服务商品化的趋势一方面的确使更多百姓更加便利地获得法律服务，拓展了业务数量和范围，但另一方面，这种商品化是否会激发对律师公共性的冲击，或者对律师事务所本身的体制和形式带来变革的压迫，都是值得深思的问题。ABS 是替代性商业结构的简称，它采用一种允许非律师人才拥有或投资建立律师事务所并提供有序法律服务的商业模式，允许律师和非律师在法律服务业务的所有权方面合作，为所有相关人员带来更大的机会。

英国在 ABS 律所的设立方面走在前列，英国律师监管机关(Solicitors Regulation Authority, SRA)截至 2012 年 10 月公布授予了四个 ABS 律所牌照。SRA 机构的独立性在于两个公会的理事不得同时担任此机构的成员，成员组成包括律师和业外人士，其组成比例为律师的人数稍多于非律师[31]，这对于 ABS 律所设立人员性质的准入上保证绝对的专业性，同时也兼顾非律师职业人员的意见。但 ABS 律所目前在中国的设立是一个充满未知和困难的进程，由于其律师和非律师人员的多种身份，以司法部、司法所对 ABS 律所进行监管的理由是否充分，争议也是非常大的。ABS 律所在中国的设立，在监管机关的归属、准入及管理体制的选择上尚需进行理论和实践的探讨。海南经济特区在这方面做了中国的首次尝试，其在《海南经济特区律师条例》第 18 条第 2 款[32]中作出放宽特殊普通合伙律师事务所合伙人准入条件的

[30] 2012 年，多伦多的 Lena Koke 和 Mark Morris 律师共同创办了 Axess Law，分布在其设置于沃尔玛超市内的品牌零售店中。店面以亮橙色为主色调，运营时间为早 10 点至晚 8 点，全年无休，提供囊括不动产转让、遗赠、遗嘱认证、授权代理及公正等在内的多种法律服务。尤其是他们的“立等可取式遗赠”服务，顾客只需要花上 99 美元，等待 45 分钟到 1 小时左右，就可以带着一份盖印生效的文书离开。其目标是在明码标价的前提下，提供一种任何人，在任何时间都可以获得的法律服务。

[31] 程幽燕：“英国的司法变迁 非律师拥有律师事务所”，载《中国律师》2013 年第 4 期，第 103 页。

[32] 《海南经济特区律师条例》第 18 条第 2 款规定：“注册会计师、注册税务师、注册造价工程师、专利代理人等其他专业人员可以成为特殊的普通合伙律师事务所的合伙人，但其出资份额和人数比例不得超过 25%，不得担任律师事务所负责人。”

规定，除了执业律师外，允许注册会计师、注册税务师、注册造价工程师、专利代理人等其他专业人士成为特殊普通合伙律师事务所的合伙人，但其人数和出资比例有一定限制，均不得超过 25%。这在拓宽法律服务相关人员执业机会的同时，也保障了律师专业判断的独立性。但在该条例出台后，并无关于专门 ABS 律所进程的讯息，其在中国的实现还有很长一段路要走。

（二）ALSP

替代性法律服务提供商（Alternative Legal Service Providers，简称 ASLP）是指除传统律师事务所外，提供法律服务的新型企业。这些企业一般不采用合伙人制度，管理人员也无需律师担任，相较于传统律师事务所，他们的公司形态较为灵活，并且可以与传统律所的非经验性的业务相竞争。[33] 替代性法律服务提供者主要分为两类，一类是跨界提供法律服务的传统企业，以四大会计师事务所的法律部门及一些投行机构为代表。由于他们在业务中本身与律所或企业有密切合作，其法律业务与律所的一些非诉业务重合度高，通过长期的积累掌握了一定的人脉和法律资源，在非诉法律业务的提供上与律所形成竞争关系；第二类是科技类的法律服务公司，其掌握了科技和数据发展的利好趋势，在法律数据的收集、储存、分析、预测等领域占据优势。相较于 ABS 而言，ALSP 的组织形式更为灵活，可以是四大会计事务所的法律部门，Legal Zoom、Elevate 等规模增长非常快的新型企业，亦或其他通过应用程序、云服务等为消费者和小企业提供法律服务的中小型企业，而 ABS 律所的侧重点在于允许非律师人员成为律所的设立者。可以看出，ABS 律所从设立到监管，均需要特定的准入规则、确定合适的监管机关、具体监管规则的限制，实践中的实现相较于 ALSP 的灵活形式而言更为困难，特别是在中国的实践中尚无先例（海南的试点并无太大进展），但在英国、澳大利亚、新加坡等国发展得很好。目前的中国显然更适合 ALSP 和 ABS 结构的发展模式，ABS 律所的发展尚待挖掘。

（三）法律公司：综合性泛法律服务提供商

法律公司相对于 ALSP 而言，更像是一个综合性泛法律服务提供商，前者囊括了 ALSP 没有的财会、IT、人力资源管理等涉法服务事项，能够提供可灵活派驻的律师，例如临时顶替企业法律部门休产假员工的空缺，或者为初创企业提供一些符合其发展阶段需求的法律服务[34]。另外从提供法律服务这一本质上来看，法律公司和律所之间并无差别，只不过在于提供法律服务的种类和方式存在差异，前者的范围会更大。所以法律公司并不希望给其冠以替代性法律服务提供商的称号，替代似乎隐含着另一种正当的方式（即传统律所），而其他方式都是存有争议和未知的。法律公司这类综合性泛法律服务提供商其实拥有广阔的市场，特别是在

---

[33] 理脉：“ALSP 风头正劲，法律服务市场路在何方？”，载 https://zhuanlan.zhihu.com/p/93800774，访问日期：2021 年 11 月 5 日。

[34] 智合东方：“法律公司（Law Company），法律行业的下一个时代？”，载 https://baijiahao.baidu.com/s?id=1622617017600389543&wfr=spider&for=pc，访问日期：2021 年 11 月 6 日。

涉法的人力资源提供层面，通过数据快速匹配符合的人力目标，弥补法律专业岗位和涉法市场的空缺。